

Maria Eugenia Mera · Manuel Morán

Measurement error covariance matrix estimation of noisy and dynamically coupled time series.

Received: date / Accepted: date

Abstract We estimate the covariance matrix of the errors in several dynamically coupled time series corrupted by measurement errors. We say that several scalar time series are dynamically coupled if they record the values of measurements of the state variables of the same smooth dynamical system. The estimation of the covariance matrix of the errors is made using a noise reduction algorithm that efficiently exploits the information contained jointly in the dynamically coupled noisy time series. The method is particularly powerful for short length time series with high uncertainties.

Keywords Measurement error models · Noise reduction · Error covariance matrix estimation · Dynamical coupling · Local projection methods

PACS 05.45.Tp · 05.10.-a · 05.45.Pq · 05.40.Ca

M. Eugenia Mera (corresponding author)

Dpto. de Fundamentos del Análisis Económico I. Universidad Complutense, 28223 Madrid. Spain

E-mail: mera@ccee.ucm.es

Manuel Morán

Dpto. de Fundamentos del Análisis Económico I. Universidad Complutense, 28223 Madrid. Spain

E-mail: mmoranca@ccee.ucm.es

1 Introduction

In many experimental situations the values of more than one feature of the phenomenon under study can be recorded as time varies. The resulting time series often contain measurement errors. Typically the covariance matrix of the corresponding multivariate measurement errors is not a multiple of the identity matrix. That is, the univariate time series have different levels of uncertainty and the corresponding measurement errors may be correlated [1]. One reason for the possible heteroskedascity and for the correlations is that in different experiments the features recorded and the devices used to obtain the data may be different. Other reasons are that those who record the data and the procedures they use vary, some characteristics are more difficult to register than others, and that the errors made in the registration of some of the features can affect the registration of other features of the phenomenon.

Many of the techniques for nonlinear time series analysis become ineffective even in the presence of low levels of measurement noise [2,3]. Medium or high noise levels suggest using a noise reduction algorithm prior to processing the observed time series, or using a convenient adaptation of standard nonlinear time series analysis techniques capable of handling noisy time series. On the other hand, since the performance of the different algorithms for estimating embedding dimension, for forecasting, or for estimating some invariant parameters such as the correlation dimension, entropy, or Lyapunov exponents often depends on the noise level [2–5], an accurate estimation of the different noise levels in the observed time series helps in selecting the method most appropriate for the observed data, for deciding what of the available time series must be used in such an estimation, and for evaluating the results obtained by the analysis.

The problem of estimating the measurement noise level in a scalar time series has attracted the interest of many researchers. The methods used can be classified as those based on correlation integrals [5–20]; those based on noise reduction algorithms [21,22]; and those based on other ideas such as logarithmic displacement [23,24], Bayesian methods [25], the gamma test [26], close neighborhoods [27], or time prediction [28].

As far as we know all the research on noise level estimation [5–28] has treated only the case of univariate time series. Even when there exist several time series, the estimation of the noise level in each scalar time series is made analyzing this time series by itself. This approach means ignoring both the information on the noise level in a given series that may be contained in the remaining time series, and possible correlations between the corresponding measurement errors.

The main contribution of this paper is that we treat the multivariate case, giving an estimate of the noise level of each of the scalar time series, and also of the correlations between the corresponding errors. A good estimate of the covariance matrix of the errors facilitates improvements in the performance of the forecasting algorithms based on local linear estimates of the unknown dynamics, that are of the most relevant methods in the forecasting literature, and obtaining confidence intervals for such predictions (see [29] for different forecasting methods). The parameters of the local linear models are estimated using the least squares method. However since, in this problem, the values of both the dependent and independent variables are subject to measurement errors, the least squares method gives biased and inconsistent estimates of the parameters of the linear models, and these models are those used in the forecasting scheme. The statistical theory which treats the problem of measurement errors in both the dependent and independent variables [1, 30–32] is called *measurement error models theory*. It provides unbiased and consistent estimates of the parameters of the linear models, which are also the maximum likelihood estimates if the errors are Gaussian. However, in order to use this theory to estimate the parameters of the linear models, it is necessary to know the covariance matrix of the errors or to have already an unbiased estimate of it [30].

We estimate the covariance matrix of the errors using the noise reduction algorithm designed for multivariate time series first proposed in [33], and improved in [34] (see [35–37] for surveys of noise reduction methods). The multivariate time series that we use as initial time series for the algorithm has among its coordinates each of the available scalar time series. The aim of any noise reduction algorithm is to isolate the deterministic component of the signal, obtaining a cleaned time series with dynamical, geometric, and statistical properties as similar as possible to those of the unknown clean time series. The difference between the initial time series and the cleaned time series is taken as the estimate of the unknown multivariate error term, and its sample covariance matrix is taken as the estimate of the covariance matrix of the unknown errors.

Our method is effective when all the scalar time series are the time ordered values of real functions (observables) of the state variables of the same smooth dynamical system. When this property holds we say that these time series are *dynamically coupled* [38] (see Ref. [39] for other approaches to coupled systems in the setting of noisy time series). We show that this method based on noise reduction of dynamically coupled time series gives good estimates of all the terms of the covariance matrix of the errors even for short length and highly corrupted time series.

That our method for the estimation of the covariance matrix of the measurement errors behaves well even for short length time series makes it possible to apply it to time series coming from the social sciences, where many of the methods used in nonlinear time series analysis have not yet been applied because of the short lengths of the available time series.

The paper is organized as follows. In Section 2 we describe briefly our noise reduction algorithm. Section 3 presents the results for the covariance matrix estimation of the error term in dynamically coupled time series generated by the logistic, Hénon, and Lorenz dynamical systems, corrupted by Gaussian noise. Finally, the conclusions are in Section 4.

2 The algorithm

2.1 The signal and the noise

We assume that we have d univariate noisy time series $X^j := \{X_i^j, i = 1, \dots, N\}$, $j = 1, \dots, d$, where

$$X_i^j = s_i^j + e_i^j; \quad i = 1, \dots, N; \quad j = 1, \dots, d. \quad (1)$$

Here $\mathbf{s}_i := (s_i^1, \dots, s_i^d)$ and $\mathbf{e}_i := (e_i^1, \dots, e_i^d)$ are, respectively, the unknown deterministic and stochastic parts of the multivariate time series $\{\mathbf{X}_i := (X_i^1, \dots, X_i^d), i = 1, \dots, N\}$. The time series $\{\mathbf{e}_i, i = 1, \dots, N\} \subset \mathbb{R}^d$ is the realization of an independent and identically distributed (*i.i.d.*) multivariate stochastic process with zero mean and unknown covariance matrix Σ . The deterministic time series $\{\mathbf{s}_i, i = 1, \dots, N\} \subset \mathbb{R}^d$ is obtained from the observation of the first N points $\{\mathbf{y}_i, i = 1, \dots, N\} \subset \mathbb{R}^m$ of an orbit of a smooth chaotic dynamical system,

$$\mathbf{s}_i = \mathbf{h}(\mathbf{y}_i) \text{ and } \mathbf{y}_{i+1} = \mathbf{f}(\mathbf{y}_i),$$

where $\mathbf{h} : M \subset \mathbb{R}^m \rightarrow \mathbb{R}^d$ and $\mathbf{f} : M \rightarrow M$ are unknown smooth functions corresponding to the observable and the evolution law respectively, and m is the unknown number of state variables of the dynamical system.

We measure the unknown noise level in the j th observed time series using the noise-to-signal standard deviation percentage (for brevity we will call it noise-to-signal level)

$$NSR^j := 100 \frac{\widehat{\sigma}(e^j)}{\widehat{\sigma}(s^j)}, j = 1, \dots, d,$$

where $\widehat{\sigma}(e^j)$ and $\widehat{\sigma}(s^j)$ denote the sample standard deviations of $e^j := \{e_i^j, i = 1, \dots, N\}$ and $s^j := \{s_i^j, i = 1, \dots, N\}$, respectively.

The multivariate time series that the algorithm takes as initial time series is $\mathbf{X}^* := \{\mathbf{X}_i^* = D(\mathbf{X}_i), i = 1, \dots, N\}$, where D is either the identity map or $D(\mathbf{X}_i) = (\mathbf{X}_i, \Delta(\mathbf{X}_i))$ for some delay operator Δ . Here are three examples for which D is not the identity map and $d = 2$. If $\Delta(\mathbf{X}_i) = \mathbf{X}_{i-1}$, then $\mathbf{X}^* = \{(X_i^1, X_i^2, X_{i-1}^1, X_{i-1}^2), i = 2, \dots, N\} \subset \mathbb{R}^4$; if $\Delta(\mathbf{X}_i) = (X_{i-1}^1, X_{i-2}^1)$, then $\mathbf{X}^* = \{(X_i^1, X_i^2, X_{i-1}^1, X_{i-2}^1), i = 2, \dots, N\} \subset \mathbb{R}^4$; and if $\Delta(\mathbf{X}_i) = (X_{i-1}^1, X_{i-1}^2, X_{i-2}^1, X_{i-2}^2)$, then $\mathbf{X}^* = \{(X_i^1, X_i^2, X_{i-1}^1, X_{i-1}^2, X_{i-2}^1, X_{i-2}^2), i = 2, \dots, N\} \subset \mathbb{R}^5$.

The dimension $d^* \geq d$ of the initial time series \mathbf{X}^* is chosen so as to guarantee a good reconstruction of the embedded dynamics (see the next subsection). The choice of what scalar time series should be delayed may depend on a prior estimate of the corresponding noise levels obtained taking $D(\mathbf{X}_i) = \mathbf{X}_i$ as the initial time series. In view of these estimates of the respective noise levels, it seems reasonable to take the time delays of those scalar time series corresponding to the lowest noise levels.

The aim of the noise reduction procedure is the separation of the deterministic component $\{\mathbf{s}_i, i = 1, \dots, N\}$ of $\{\mathbf{X}_i, i = 1, \dots, N\}$ from the observational noise $\{\mathbf{e}_i, i = 1, \dots, N\}$. Let $\{\hat{\mathbf{s}}_i, i = 1, \dots, N\}$ be the estimation of $\{\mathbf{s}_i, i = 1, \dots, N\}$ given by the noise reduction algorithm. The estimation of $\{\mathbf{e}_i, i = 1, \dots, N\}$ is $\{\hat{\mathbf{e}}_i := \mathbf{X}_i - \hat{\mathbf{s}}_i, i = 1, \dots, N\}$, and the sample covariance matrix $\hat{\Sigma}$ of $\{\hat{\mathbf{e}}_i, i = 1, \dots, N\}$ will be the estimate of Σ . Then the estimate of the noise-to-signal level in the j th observed time series is

$$\widehat{NSR}^j := 100 \frac{\widehat{\sigma}(\hat{e}^j)}{\widehat{\sigma}(\hat{s}^j)}, j = 1, \dots, d,$$

and the estimate of the correlation between the k th and the j th coordinates of the error is

$$\hat{R}_{kj} := \frac{\hat{\Sigma}_{k,j}}{\sqrt{\hat{\Sigma}_{k,k} \hat{\Sigma}_{j,j}}}, k = 1, \dots, d; j = 1, \dots, d,$$

where $\hat{\Sigma}_{k,j}$ is the entry in the k th row and j th column of the matrix $\hat{\Sigma}$.

2.2 The cleaning procedure

We recall now the main ideas of our noise reduction algorithm. The details of its implementation are given in Ref. [34]. We first take a three-delay embedding $\mathbf{Z} := \{\mathbf{Z}_i := (\mathbf{X}_{i-1}^*, \mathbf{X}_i^*, \mathbf{X}_{i+1}^*), i = 2, \dots, N-1\} \subset \mathbb{R}^{3d^*}$ of the initial time series \mathbf{X}^* . The dimension $d^* \geq d$ of the vectors in \mathbf{X}^* is chosen so that the time series \mathbf{Z} provides a good reconstruction of the underlying dynamics. This is guaranteed if $3d^* > 2m$ (see Refs. [40, 41]). In the examples reported below we have taken larger values of d^* in the cases of high noise levels in order to reduce

the number of false neighbors (points which are close only due to the noise). In order to simplify the notation we assume that $d^* = d$, that is $D(\mathbf{X}_i) = \mathbf{X}_i$ and therefore $\mathbf{X}_i^* = \mathbf{X}_i$, $i = 1, \dots, N$, but the method is the same in the general case.

As a result of the noise, $\{\mathbf{Z}_i\}$ does not lie in an m -dimensional submanifold of \mathbb{R}^{3d} . Let U_i be a neighborhood of \mathbf{Z}_i , $i = 2, \dots, N-1$. The noise is partially reduced through orthogonal projections of the data points within each U_i onto the m -dimensional linear subspace which is the closest, with respect to a given metric, to the data points in U_i . Our goal is to exploit in an optimal way the information contained in all the scalar time series. Since the levels of uncertainty of the various scalar time series may be very different, we use the metric induced by an estimate $\widehat{\Sigma}_3$ of the covariance matrix Σ_3 of the unknown errors contained in $\{\mathbf{Z}_i, i = 2, \dots, N-1\}$. This distance gives less weight to the coordinates (scalar time series) recorded with lower precision.

The use of the metric induced by $\widehat{\Sigma}_3$ might seem surprising because our goal is the estimation of the covariance matrix Σ of the errors in $\{\mathbf{X}_i, i = 1, \dots, N\}$. Since the noise reduction is an iterative scheme, $\widehat{\Sigma}_3$ can be obtained through a successive refinement process. The initial time series $\{\mathbf{X}_i^{(k)}, i = 1, \dots, N\} \subset \mathbb{R}^d$ at a given iteration k is the output time series $\{\mathbf{s}_i^{(k-1)}, i = 1, \dots, N\}$ of the previous iteration of the algorithm. At each iteration k there is needed an estimate $\widehat{\Sigma}_3^{(k)}$ of the covariance matrix $\Sigma_3^{(k)}$ of the remaining errors contained in the three-delay embedding $\{\mathbf{Z}_i^{(k)}, i = 2, \dots, N-1\}$ of the initial time series $\{\mathbf{X}_i^{(k)}, i = 1, \dots, N\}$ at this iteration. For $k \geq 2$, we take as $\widehat{\Sigma}_3^{(k)}$ the sample covariance matrix of the corrections $\{\mathbf{Z}_i^{(k)} - \mathbf{Z}_i^{(k-1)}, i = 2, \dots, N-1\}$ made by the algorithm in the previous iteration. However, there remains the issue of estimating the covariance matrix to be used in the first iteration of the algorithm. If we do not have any a priori information about Σ_3 , we take the identity matrix as the estimate in the first iteration. Then we run several iterations of the algorithm (eight iterations suffices in all our experiments), and take as an a posteriori estimate of Σ_3 the sample covariance matrix of the total corrections made by the algorithm in these eight iterations. This estimate $\widehat{\Sigma}_3$ is then used in the first iteration of the full noise reduction scheme.

Let $\widehat{\Sigma}_3^{(k)}$ be the estimate of the covariance matrix of the remaining errors in the three-delay embedding $\{\mathbf{Z}_i^{(k)}, i = 2, \dots, N-1\}$ of the initial time series $\{\mathbf{X}_i^{(k)}, i = 1, \dots, N\}$ at iteration k (below, in order to simplify the notation we omit the superscript k corresponding to the iteration). For each point \mathbf{Z}_i , we take a neighborhood U_i of \mathbf{Z}_i , and the center of mass $\langle \mathbf{Z}_i \rangle$ of the data points in U_i . The best m -dimensional linear subspace \mathbf{T}_i for

the vectors $\{\mathbf{Z}_j - \langle \mathbf{Z}_i \rangle : \mathbf{Z}_j \in U_i\}$ is the solution of

$$\min_{\mathbf{T} \in \mathcal{L}_m} \sum_{j: \mathbf{Z}_j \in U_i} \left(d_{\hat{\Sigma}_3, \mathbf{T}}(\mathbf{Z}_j - \langle \mathbf{Z}_i \rangle) \right)^2,$$

where \mathcal{L}_m is the set of all m -dimensional linear subspaces of \mathbb{R}^{3d} , $d_{\hat{\Sigma}_3, \mathbf{T}}(\mathbf{Z}) := \min_{\mathbf{u} \in \mathbf{T}} d_{\hat{\Sigma}_3}(\mathbf{Z}, \mathbf{u}) = d_{\hat{\Sigma}_3}(\mathbf{Z}, \mathbf{P}_{\mathbf{T}}\mathbf{Z})$ is the distance of the vector \mathbf{Z} to the linear subspace $\mathbf{T} \in \mathcal{L}_m$ in the metric induced by $\hat{\Sigma}_3$, $\mathbf{P}_{\mathbf{T}}\mathbf{Z}$ is the orthogonal projection of \mathbf{Z} onto \mathbf{T} , and $d_{\hat{\Sigma}_3}(\mathbf{Z}, \mathbf{u}) := \left((\mathbf{Z} - \mathbf{u})^t \hat{\Sigma}_3^{-1} (\mathbf{Z} - \mathbf{u}) \right)^{1/2}$ is the distance between the vectors \mathbf{Z} and \mathbf{u} in the metric induced by $\hat{\Sigma}_3$. This is an optimization problem with an explicit solution easily found numerically [34].

The noise in $\mathbf{Z}_j \in U_i$ is partially removed by replacing \mathbf{Z}_j by $\hat{\mathbf{Z}}_{j(i)} := \langle \mathbf{Z}_i \rangle + \mathbf{P}_{\mathbf{T}_i}(\mathbf{Z}_j - \langle \mathbf{Z}_i \rangle)$. Since \mathbf{Z}_j can belong to many neighborhoods U_i , we have an estimate $\hat{\mathbf{Z}}_{j(i)}$ for each such neighborhood. The final estimate $\hat{\mathbf{Z}}_j := (\hat{\mathbf{s}}_{j-1}, \hat{\mathbf{s}}_j, \hat{\mathbf{s}}_{j+1})$ of $(\mathbf{s}_{j-1}, \mathbf{s}_j, \mathbf{s}_{j+1}) \in \mathbb{R}^{3d}$ obtained at a given iteration is the average of all of these estimates. This gives a $3d$ -dimensional time series and therefore three different estimates for each \mathbf{s}_i , $i = 2, \dots, N-1$: those given respectively by the final, intermediate, and initial d coordinates of $\hat{\mathbf{Z}}_{i-1}$, $\hat{\mathbf{Z}}_i$, and $\hat{\mathbf{Z}}_{i+1}$. We take as the estimate of \mathbf{s}_i a weighted average of these three estimates. The weights are the same for all the points and inversely proportional to the sample variances of the corrections $\varepsilon_i := (\mathbf{X}_{i-1}, \mathbf{X}_i, \mathbf{X}_{i+1}) - (\hat{\mathbf{s}}_{i-1}, \hat{\mathbf{s}}_i, \hat{\mathbf{s}}_{i+1})$, $i = 2, \dots, N-1$.

The effectiveness of the noise reduction scheme means that noise levels decrease as the iterations advance. In our algorithm the shapes and sizes of the neighborhoods are automatically determined by the remaining level of uncertainty in each coordinate. They are defined as $U_i = \{\mathbf{Z}_j : d_{\hat{\Sigma}_3}(\mathbf{Z}_j, \mathbf{Z}_i) < r\}$ where r is taken as small as possible subject to the requirement that, with a given confidence level, the relevant information is contained in the neighborhoods [34]. The neighborhoods will be ellipsoids rather than simply spheres, with semi-axial lengths related to the uncertainties in the corresponding coordinates. A clear indication that further iterations of the algorithm will not reduce the remaining noise is the stabilization of the mean number of points in the neighborhoods, and this is our stopping criterion.

2.3 Alternative methods based on local orthogonal projections

Noise reduction via orthogonal projections of the data in neighborhoods onto linear subspaces is also the basis of other noise reduction algorithms designed for scalar [21, 22, 35, 42–47], and multivariate [48] time series.

Since these algorithms work with delay vectors obtained from scalar time series, the levels of uncertainty in all the coordinates of such a delay vector are the same. In many of these algorithms the Euclidean metric is used for estimating the linear subspaces onto which the data in the neighborhoods are projected [21, 43, 44, 47]. Another metric which gives good results for scalar time series is the one proposed in Ref. [35]. This metric focuses the noise reduction on the most stable middle coordinates of the delay vectors. However, since these algorithms are designed for scalar time series they do not take advantage of the information about the levels of uncertainty that several dynamically coupled scalar time series may contain jointly, nor do they permit estimation of the correlations between the corresponding coordinates of the error.

For the purpose of estimating the covariance matrix of the errors contained in a multivariate time series there can be used any noise reduction algorithm designed for multivariate time series (for instance that proposed in [48] and implemented in the *ghkss* routine of the TISEAN library [49]). However, as far as we know, the algorithm first proposed in [33] and later improved in [34] is the only one designed for multivariate time series that takes into account either the possible differences between the levels of uncertainty of different coordinates of the time series or the correlations between the coordinates of the error. The performance of the noise reduction algorithm depends strongly on the metric considered, because the metric determines which linear subspace \mathbf{T}_i is closest to the data points in the neighborhood U_i , what is the orthogonal projection $P_{\mathbf{T}_i}(\mathbf{Z}_j - \langle \mathbf{Z}_i \rangle)$ of a vector $\mathbf{Z}_j - \langle \mathbf{Z}_i \rangle$ onto the linear subspace \mathbf{T}_i , and, in our algorithm, also what are the shapes and sizes of the neighborhoods U_i . The metric also determines the statistical properties of the output time series. The Euclidean metric gives biased estimators [50–52], whereas the metric induced by the estimated covariance matrix of the errors gives unbiased and consistent estimates of the true values for the case of linear models and, for Gaussian errors, these estimates are those of maximum likelihood [30]. For all these reasons we think that it is best to use an algorithm based on the metric induced by $\hat{\Sigma}_3$, especially in the cases where the time series have short lengths, where the levels of uncertainty in each of the scalar time series are very different, or where the coordinates of the multidimensional error are highly correlated.

3 Numerical results

In this section we show the results obtained for time series generated by logistic, Hénon, and Lorenz dynamics corrupted by Gaussian measurement noise.

The logistic dynamics is generated by the map

$$y(k+1) = 4y(k)(1-y(k)).$$

The Hénon dynamical system is generated by the mapping

$$y_1(k+1) = 1 - ay_1(k)^2 + y_2(k), \quad y_2(k+1) = by_1(k),$$

where the parameter values are taken to be $a = 1.4$ and $b = 0.3$. The Lorenz dynamical system is defined by

$$\dot{y}_1 = \sigma(y_2 - y_1), \quad \dot{y}_2 = y_1(R - y_3) - y_2, \quad \dot{y}_3 = y_1y_2 - \beta y_3,$$

where the parameter values are taken to be $\sigma = 16$, $R = 45.92$, and $\beta = 4$. These equations are integrated using a fourth order Runge-Kutta algorithm with an integration step of $\Delta t = 0.001$. The sampling time we consider is $30\Delta t = 0.03$.

For short length time series the results depend strongly on the realization of the error term, and also on the clean time series considered. For this reason we repeat each experiment several times for different clean time series and realizations of the error term. We generate, using one of the three dynamical systems mentioned above, a long clean time series of 50000 data points. This clean time series is observed using different scalar observables h^j , $j = 1, \dots, d$. The resulting d -dimensional time series of 50000 data points is corrupted with an i.i.d. d -dimensional Gaussian noise with a given covariance matrix. We use the subroutine DRNMVN of the IMSL libraries [53] to generate this Gaussian noise with a given covariance matrix. Finally we split each of these d time series into L time series each having $N = 50000/L$ data points. We have considered times series having lengths $N \in \{250, 1000, 5000\}$, in order to study the behavior of the algorithm for time series of very short, medium, and large lengths.

In each of these L experiments we estimate the sample covariance matrix of the errors $\hat{\Sigma}$, giving as the result the sample mean $\langle \hat{\Sigma} \rangle$ of the sample covariances matrices obtained from these L experiments. For comparison we also include the sample mean $\langle \Sigma \rangle$ of the sample covariance matrices corresponding to the true errors. We also report the $d \times d$ matrices $\sigma_{\hat{\Sigma}}$ and σ_{Σ} whose entries are the sample standard deviations $\sigma_{\hat{\Sigma}_{i,j}}$ and $\sigma_{\Sigma_{i,j}}$, $i = 1, \dots, d$, $j = 1, \dots, d$, of the elements of these covariance matrices. Thus for $N = 250$, the mean matrices $\langle \hat{\Sigma} \rangle$ and $\langle \Sigma \rangle$ are calculated using 200 matrices, whereas for $N = 5000$ they are calculated using 10 matrices.

In order to evaluate better the relative errors made in estimating the elements of the covariance matrix, we also include in the results reported below the mean values $\langle \widehat{NSR}^j \rangle, j = 1, \dots, d$, of the noise-to-signal levels and the mean values $\langle \widehat{R}_{ij} \rangle, i = 1, \dots, d, j = 1, \dots, d$, of the correlations between the coordinates of the estimated errors, together with their respective sample standard deviations $\sigma_{\widehat{NSR}^j}, j = 1, \dots, d$, and $\sigma_{\widehat{R}_{i,j}}, i = 1, \dots, d, j = 1, \dots, d$. In the tables reported below we denote by $\langle \widehat{\Gamma} \rangle$ the matrix whose diagonal entries are $\langle \widehat{NSR}^j \rangle, j = 1, \dots, d$, and whose upper triangular entries (i, j) are $\langle \widehat{R}_{ij} \rangle, i = 1, \dots, d, j = i + 1, \dots, d$. The matrix containing the corresponding sample standard deviations of the entries in $\langle \widehat{\Gamma} \rangle$ is denoted $\sigma_{\widehat{\Gamma}}$. The analogous matrices for the true d -dimensional error are denoted $\langle \Gamma \rangle$ and σ_{Γ} respectively.

3.1 One-dimensional discrete dynamics and small noise-to-signal ratios

In this subsection we present the results obtained by applying our algorithm to three dynamically coupled noisy time series. These three time series are generated using the logistic map and the scalar observables $h^1(y) = \sqrt{4+5y}$, $h^2(y) = (2+y)^2$, and $h^3(y) = \frac{\ln(y+1)+2}{\ln 2}$ respectively. The three-dimensional error term has a Gaussian distribution, with low noise-to-signal levels of 4%, 5%, and 6% respectively, and correlations of $R_{12} = 0.8$, $R_{13} = 0.2$, and $R_{23} = -0.2$, that is $\Gamma = \begin{pmatrix} 4 & 0.8 & 0.2 \\ & 5 & -0.2 \\ & & 6 \end{pmatrix}$.

The aims of this experiment are to analyze if the algorithm detects the subtle differences between the three coordinates of the measurement error with respect to the noise levels, and to see if it provides reasonable estimates of the correlations between the coordinates of the error.

The results presented in Table 1 correspond to the initial time series $\{\mathbf{X}_i = (X_i^1, X_i^2, X_i^3), i = 1, \dots, N\}$. The dimension m of the linear subspaces where the data are projected is taken to be $m = 1$.

Table 1(a) displays, as a function of the length N of the time series, the matrices $\langle \Gamma \rangle$ and σ_{Γ} on its left side, and their respective estimates $\langle \widehat{\Gamma} \rangle$ and $\sigma_{\widehat{\Gamma}}$ on its right. The matrices $\langle \Sigma \rangle$, σ_{Σ} , $\langle \widehat{\Sigma} \rangle$ and $\sigma_{\widehat{\Sigma}}$ are shown in Table 1(b). We can see that the algorithm detects the subtle differences between these three low levels of noise even for time series of only 250 data points. The estimates of the mean noise-to-signal levels and also the estimated correlations are good even for very short time series, and they improve for longer time series. The empirical standard deviations $\sigma_{\widehat{NSR}^j}, j = 1, \dots, 3$, and $\sigma_{\widehat{R}_{i,j}}, i = 1, \dots, 3, j = 1, \dots, 3$, corresponding to the estimated values decrease with the length of the time series and they are only slightly higher than the sample standard deviations $\sigma_{NSR^j}, j = 1, \dots, 3$, and $\sigma_{R_{i,j}}, i = 1, \dots, 3, j = 1, \dots, 3$, obtained with the true error values.

For long time series the relative error made in the estimation of the entries of $\langle \Gamma \rangle$ is very small, less than 0.5% in all the entries with the exception of R_{23} , for which the relative error is about 1.14%.

The estimates of the entries in $\langle \Sigma \rangle$ and their corresponding standard deviations are also very good (see Table 1(b)), even for time series of only 250 data points.

3.2 Two-dimensional discrete dynamics and high amplitude noises

In Table 2(a-b) we present the results obtained for highly corrupted time series. We have considered two dynamically coupled noisy time series obtained from the observation, using the observables $h^1(y_1, y_2) = \frac{50y_2}{5+y_1}$ and $h^2(y_1, y_2) = 4y_1^2 + y_2$, of an orbit generated by the Hénon map. The two dimensional error has a Gaussian distribution, with high noise-to-signal levels of 50% and 25%, respectively, and a correlation of $R_{12} = 0.75$.

The aim of this experiment is to analyze if the algorithm estimates well the noise levels and the correlations between the error terms for highly corrupted time series. The results presented in Table 2 correspond to the initial time series $\mathbf{X}^* = \{(\mathbf{X}_i, \mathbf{X}_{i-1}, \mathbf{X}_{i-2}), i = 3, \dots, N+2\}$, where $\mathbf{X}_i = (X_i^1, X_i^2)$, and taking $m = 2$ as the dimension of the linear subspaces where the data points are projected.

The results displayed in Table 2 show that the algorithm provides good estimates of the entries of Σ and Γ even for very short length time series, and the results improve when the lengths of the time series increase, mainly because the mean sample standard deviations of all the estimates decrease quickly when the length of the time series increases.

In the interest of brevity we do not present the results taking $\mathbf{X}^* = \{(\mathbf{X}_i, \mathbf{X}_{i-1}), i = 2, \dots, N+1\}$ or $\mathbf{X}^* = \{(\mathbf{X}_i, D(\mathbf{X}_i), i = 3, \dots, N+2\}$ with $D(\mathbf{X}_i) = (X_{i-1}^1, X_{i-1}^2, X_{i-2}^1)$ as the initial time series. Although the results are also good, those for higher dimensional initial time series are better, because the effects of false neighbors are smaller.

3.3 Three-dimensional continuous dynamics and moderate amplitude noises

The results contained in Table 3 (a-b) correspond to three dynamically coupled time series generated using the Lorenz dynamics and the three scalar observables $h^1(y_1, y_2, y_3) = y_1 + 5y_2^2 + y_3$, $h^2(y_1, y_2, y_3) = 2y_2 + y_3^2$, and $h^3(y_1, y_2, y_3) = 2y_1 + y_3^2$, respectively. The three dimensional error term has a Gaussian distribution, with moderate noise-to-signal levels of 5%, 10%, and 20%, respectively. The correlations between the coordinates

of the error are $R_{12} = 0.1$, $R_{13} = 0$, and $R_{23} = -0.1$. The first aim of this experiment is to analyze the behavior of the algorithm for time series generated by three-dimensional dynamics and corrupted by noises of moderate amplitude. The second aim is to analyze if the algorithm distinguishes between a null correlation and a small correlation (positive or negative).

In order to guarantee that the initial time series provides a good reconstruction of the underlying dynamics and to reduce the problems caused by false neighbors, we take as the initial time series $\mathbf{X}^* = \{(\mathbf{X}_i, \mathbf{X}_{i-1}, \mathbf{X}_{i-2}), i = 3, \dots, N+2\}$, where $\mathbf{X}_i = (X_i^1, X_i^2, X_i^3)$. The dimension m of the linear subspaces where the data are projected is taken to be $m = 3$.

The estimates of the mean noise-to-signal levels are reasonable even for short length time series. However, in order to determine with confidence the signs of such small correlations we need at least $N = 1000$ data points.

In the interest of brevity we do not present the results obtained taking $\mathbf{X}^* = \{(\mathbf{X}_i, \mathbf{X}_{i-1}), i = 2, \dots, N+1\}$ as the initial time series, but these results are quite similar to those displayed in Table 3, showing the robustness of the output of the algorithm with respect to the dimension of the initial time series, provided the chosen dimension allows a faithful reconstruction of the underlying dynamics, and that for such a dimension the problem of false neighbors is not very important (which obviously depends on the noise levels).

A comparison of the empirical distribution of the estimated errors as a function of the lengths N of the time series used in the noise reduction algorithm is shown in Figure 1. We have plotted only the histograms corresponding to the relative frequencies of the third component of the estimated error term, which is the one with the highest variance when compared with the variance of the corresponding coordinate of the clean signal. The results for the other two coordinates of the error are analogous. In Figure 1(a) the estimated errors are obtained by splitting the noisy time series having 50000 data points into 200 noisy time series of 250 data points each (short length time series), and processing each of these 200 noisy time series separately. In Figure 1(b) the estimated errors are obtained by splitting the noisy time series having 50000 data points into 10 time series of 5000 data points. Both figures exhibit also the histogram corresponding to the third coordinate of the true errors.

We can see that the empirical distribution of the estimated errors is close to the distribution of the true errors even for the case of short length time series. These figures suggest strongly the convergence of the

distribution of the estimated errors to the distribution of the true errors when the length N of the time series used to estimate the errors increases.

4 Conclusions

We propose a method which allows the estimation of all the entries of the covariance matrix of the additive multivariate error with which a deterministic multivariate time series has been corrupted. Our method improves upon the existing methods in the literature. Because these only treat noise level estimation for the case of scalar noisy time series, they do not take advantage of the information about the error that a joint treatment of all the available information may provide, nor do they estimate the correlations between the coordinates of the error in the case that several scalar time series are available.

An efficient method for estimating the covariance matrix of the measurement errors may be essential for improving the behavior of forecasting algorithms, for deciding what methods or what of the available scalar time series must be used for analyzing the properties of the underlying dynamics, or for a better evaluation of the reliability of the results obtained by the application of a given method to a noisy time series.

Although the estimation of noise levels in scalar time series by a noise reduction algorithm has been proposed before (see [21,22]), the noise reduction algorithms used depend on many parameters such as the embedding dimension, size of the neighborhoods, number of iterations, etc., making their use for noise level estimation difficult especially for large noise-to-signal standard deviation ratios. The noise reduction algorithm we use does not depend on so many parameters. It incorporates devices to determine the sizes of the neighborhoods used in each iteration, and also when the algorithm should terminate. The only parameters that remain free are the dimension of the initial time series and the dimension of the linear subspaces where the data are projected. However we have checked that the results do not change substantially when the embedding dimension is changed, provided that it is such that the underlying dynamics can be faithfully reconstructed, and the problem of false neighbors can be controlled. Although the results are better when the dimension p of the linear subspaces where the data are projected coincides with the dimension m of the space where the dynamics is defined, the results are also reasonable if $p > m$ and the embedding dimension is chosen large enough.

Even if the goal is only to estimate the noise-to-signal ratio of a single scalar time series, a better estimate is obtained if this scalar time series is processed jointly with other scalar time series dynamically coupled to it and with a lower noise-to-signal standard deviation ratio.

Our method gives very good estimates of all the entries of the covariance matrix of the errors even for very short length time series and measurement noises with large noise-to-signal standard deviation ratios. Our experiments also suggest strongly that the distribution of the estimated errors converges to the distribution of the true error term when the length of the time series increases, even in the case when the noise-to-signal standard deviation ratio is very high.

Acknowledgements This research was supported by the grant MTM2009-12672 given by the Spanish government's Ministerio de Ciencia e Innovación.

References

1. Buonaccorsi, J. P.: Measurement Error Models, Methods, and Applications. Chapman & Hall/CRC, Boca Raton (2010)
2. Kantz, H., Schreiber, T.: Nonlinear Time Series Analysis. Cambridge University Press, Cambridge (2004)
3. Abarbanel, H. D. I.: Analysis of Observed Chaotic Data. Springer-Verlag, New York (1997)
4. Molkov, Y. I., Mukhin, D. N., Loskutov, E. M., Feigin, A. M.: Using the minimum description length principle for global reconstruction of dynamic systems from noisy time series. Phys. Rev. E 80, 046207 (2009)
5. Diks, C.: Nonlinear time series analysis. Methods and Applications. World Scientific, Singapore (1999)
6. Smith, R.L.: Estimating dimension in noisy chaotic time series. J.R. Stat. Soc. B 54, 329-351 (1992)
7. Schreiber, T.: Determination of the noise level of chaotic time series. Phys. Rev. E 48, R13-R16 (1993)
8. Schouten, J. C., Takens, F., van den Bleek, C. M.: Estimation of the dimension of a noisy attractor. Phys. Rev. E 50, 3, 1851 (1994)
9. Diks, C.: Estimating invariants of noisy attractors. Phys. Rev. E 53, R4263R4266 (1996)
10. Schreiber, T.: Influence of Gaussian noise on the correlation exponent. Phys. Rev. E 56, 1, 274-277 (1997)
11. Oltmans, H., Verheijen, P. J. T.: Influence of noise on power-law scaling functions and an algorithm for dimension estimations. Phys. Rev. E 56, 1, 1160-1170 (1997)
12. Kugiumtzis, D.: Correction of the Correlation Dimension for Noisy Time Series. Int. J. Bifurcations Chaos 7, 6, 1283-1294 (1997)
13. Galka, A.: Topics in Nonlinear Time Series Analysis with Implications for EEC Analysis. World Scientific Publishing, Singapore (2000)
14. Yu, D., Small, M., Harrison, R.G., Diks, C.: Efficient implementation of the gaussian kernel algorithm in estimating invariants and noise level from noisy time series data. Phys. Rev. E 61, 4, 3750-3756 (2000)

15. Leontitsis, A., Pange, J., Bountis, T.: Large noise level estimation. *Int. J. Bifurcations Chaos* 13, 8, 2309-2313 (2003)
16. Urbanowicz, K., Holyst, J.A.: Noise-level estimation of time series using coarse-grained entropy. *Phys. Rev. E* 67, 046218 (2003)
17. Kaminski, T., Wendeker, M., Urbanowicz, K., Litak, G.: Combustion process in a spark ignition engine: Dynamics and noise level estimation. *Chaos* 14, 2, 461 (2004)
18. Strumik, M., Macek, W. M., Redaelli, S.: Discriminating additive from dynamical noise for chaotic time series. *Phys. Rev. E* 72, 036219 (2005)
19. Leontitsis, A.: Red Noise Estimation. *Int. J. Bifurcations Chaos* 16, 7, 2111-2117 (2006)
20. Jayawardena, A. W., Xu, P., Li, W. K.: A method of estimating the noise level in a chaotic time series. *Chaos* 18, 023115 (2008)
21. Cawley, R., Hsu, G-H.: Local-geometric-projection method for noise reduction in chaotic maps and flows. *Phys. Rev. A* 46, 6, 3057-3082 (1992)
22. Urbanowicz, K., Holyst, J.A.: Noise reduction in chaotic time series by a local projection with nonlinear constraints. *Acta Phys. Pol. B*, 35, 9, 2175-2197 (2004)
23. Cellucci, C. J., Albano, A. M., Rapp, P. E., Pittenger, R. A., Josiassen, R.C.: Detecting noise in a time series. *Chaos* 7 (3), 414-422 (1997)
24. Hu, J., Gao, J.B., White, K. D.: Estimating measurement noise in a time series by exploiting nonstationarity. *Chaos, Solitons & Fractals* 22, 807-819 (2004)
25. Heald, J. P. M., Stark, J.: Estimation of noise levels for models of chaotic dynamical systems. *Phys. Rev. Lett.* 84, 11, 2366-2369 (2000)
26. Jones A. J., Evans D., Kemp S. E.: A note on the Gamma test analysis of noisy input/output data and noisy time series. *Physica D* 229, 1-8 (2007)
27. Urbanowicz, K., Holyst, J. A.: Noise estimation by use of neighboring distances in takens space and its applications to stock market data. *Int. J. Bifurcations Chaos* 16, 6, 1865-1869 (2006)
28. Tanaka, N., Okamoto, H., Naito, M.: Estimating the amplitude of measurement noise present in chaotic time series. *Chaos* 9, 2, 436-444 (1999)
29. Soofi, A. S., Cao, L. (eds.): *Modeling and Forecasting Financial data*. Kluwer academic publishers, Boston (2002)
30. Fuller, W. A.: *Measurement error models*. Wiley Series in Probability and Statistics. John Wiley & Sons, New York (2006)
31. Cheng, C-L., Van Ness, J. W.: *Statistical Regression with Measurement Error*. Arnold, London (1999)
32. Carroll, R. J., Ruppert, D., Stefanski, L. A., Crainiceanu, C. M.: *Measurement Error in Nonlinear Models. A Modern Perspective*. Chapman & Hall/CRC, New York (2006)
33. Mera, M.E., Morán, M.: Geometric noise reduction for multivariate time series. *Chaos*, 16, 013116 (2006)
34. Mera, M.E., Morán, M.: Reduction of noise of large amplitude through adaptive neighborhoods. *Phys. Rev. E* 80, 016207 (2009)
35. Grassberger, P., Hegger, R., Kantz, H., Schaffrath, C., Schreiber, T.: On noise reduction methods for chaotic data. *Chaos* 3, 2, 127-141 (1993)

-
36. Kostelich, E. J. , Schreiber, T.: Noise reduction in chaotic time-series data: A survey of common methods. *Phys. Rev. E* 48, 1752-1763 (1993)
 37. Davies, M.: Noise reduction schemes for chaotic time series. *Physica D* 79, 174-192 (1994)
 38. Mera, M.E., Morán, M.: Noise reduction by recycling dynamically coupled time series. *Chaos*, 21, 043110 (2011)
 39. Netoff, T. I., Carroll, T. L., Pecora, L. M., Schiff, S. J.: Detecting Coupling in the Presence of Noise and Nonlinearity. In: *Handbook of Time Series Analysis*. Scheler, B., Winterhalder, M., Timmer, J. (eds.), 265-282. Wiley-VCH, Weinheim (2006)
 40. Takens, F.: Dynamical Systems and turbulence. In: Rang, D.A., Young, L.-S.(eds.). *Lecture Notes in Mathematics* 898, 366-381. Springer, Berlin, (1981)
 41. Sauer, T., Yorke, J., Casdagli, M.: Embedology. *J. Stat. Phys.*, 65, 579-616 (1991)
 42. Kern, A., Steeb, W.-H., Stoop, R.: Projective noise cleaning with dynamic neighborhood selection. *Int. J. of Mod. Phys. C*, 11, 1, 125-146 (2000)
 43. Matassini, L., Kantz, H., Holyst, J., Hegger, R.: Optimizing of recurrence plots for noise reduction. *Phys. Rev. E* 65, 021102 (2002)
 44. Leontitsis, A., Bountis, T., Pagge J.: An adaptive way for improving noise reduction using local geometric projection. *Chaos* 14, 1, 106 (2004)
 45. Johnson, M.T., Povinelli, R. J.: Generalized phase space projection for nonlinear noise reduction. *Physica D* 201, 306-317 (2005)
 46. Luo, X., Zhang, J., Small, M.: Optimal phase-space projection for noise reduction. *Phys. Rev. E* 72, 046710 (2005)
 47. Sauer, T.: A noise reduction method for signals from nonlinear systems. *Physica D* 58, 193-201 (1992)
 48. Hegger, R., Schreiber, T.: A noise reduction method for multivariate time series. *Phys. Lett. A* 170, 305-310 (1992)
 49. Hegger, R., Kantz H., Schreiber, T.: Practical implementation of nonlinear time series methods: The TISEAN package. *Chaos* 9, 413 (1999)
 50. Kostelich, E. J., Yorke, J. A.: Noise reduction: Finding the simplest dynamical system consistent with the data. *Physica D*, 41, 183-196 (1990)
 51. Jaeger, L., Kantz, H.: Unbiased reconstruction of the dynamics underlying a noisy chaotic time series. *Chaos* 6, 440 (1996)
 52. Bröcker, J., Parlitz, U.: Efficient noncausal noise reduction for deterministic time series. *Chaos* 11, 2, 319 (2001)
 53. International Mathematics and Statistics Libraries (IMSL 4.01). Visual Numerics Inc. (1997)

Fig. 1 Histogram corresponding to the relative frequencies of the true (in black) and estimated (in red) third coordinate of the error term corresponding to the experiment with the Lorenz dynamics.

The true error term has a standard deviation equal to 20% of the standard deviation of the corresponding clean signal. In (a) the estimated errors are obtained by processing separately 200 noisy time series each of 250 data points, and in (b) using 10 noisy time series each of 5000 data points.

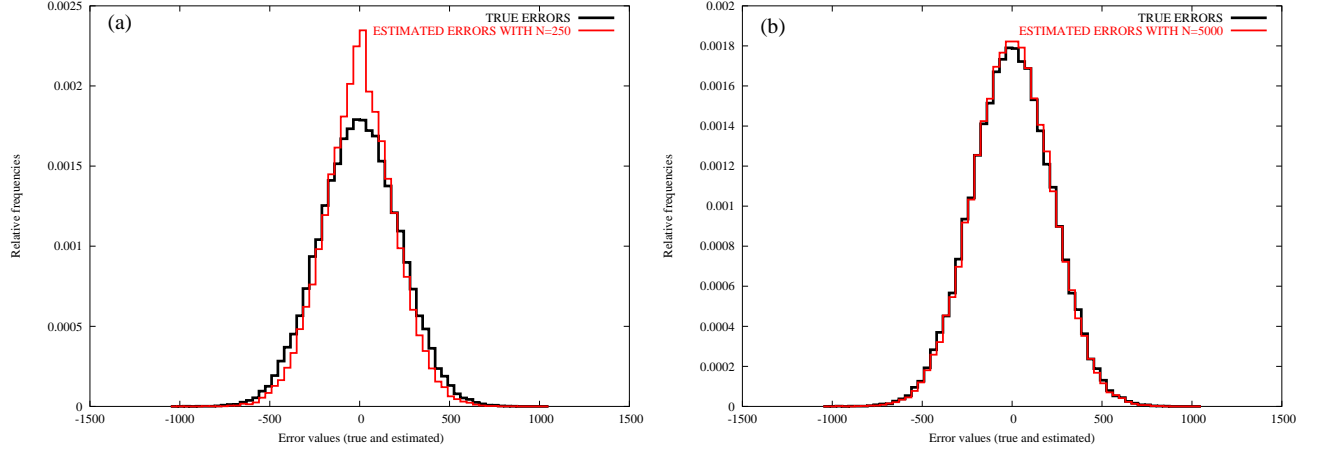


Table 1 Results using three scalar logistic time series corrupted with low level noises.

The three noisy scalar time series are obtained from the observation of an orbit of the logistic map using the observables $h^1(y) = \sqrt{4+5y}$, $h^2(y) = (2+y)^2$, and $h^3(y) = \frac{\ln(y+1)+2}{\ln 2}$, respectively, and corrupted by noises with noise-to-signal levels of 4%, 5%, and 6%, respectively, and with correlation values of $R_{12} = 0.8$, $R_{13} = 0.2$, and $R_{23} = -0.2$.

(a) True (on the left of the table) and estimated (on the right of the table) mean noise-to-signal levels, and mean correlations between the coordinates of the three dimensional error term obtained using L time series each of N data points. The mean noise-to-signal levels are in the diagonal entries of the matrices, and the mean correlations are in the upper diagonal terms of these matrices. The matrices σ_Γ and $\sigma_{\hat{\Gamma}}$ located beside the corresponding matrix $\langle \Gamma \rangle$ or $\langle \hat{\Gamma} \rangle$ contain the standard deviations.

(b) True (on the left of the table) and estimated (on the right of the table) mean covariance matrices obtained using L time series each of N data points. The corresponding standard deviations are the entries of the matrix beside the covariance matrix.

(a)

(N, L)	$\langle \Gamma \rangle \pm \sigma_\Gamma$	$\langle \hat{\Gamma} \rangle \pm \sigma_{\hat{\Gamma}}$
(250, 200)	$\begin{pmatrix} 4.0128 & 0.8003 & 0.2068 \\ & 4.9977 & -0.1927 \\ & & 6.0013 \end{pmatrix} \pm \begin{pmatrix} 0.1889 & 0.0238 & 0.0615 \\ & 0.2483 & 0.0579 \\ & & 0.2709 \end{pmatrix}$	$\begin{pmatrix} 3.7683 & 0.7869 & 0.1977 \\ & 4.6884 & -0.2165 \\ & & 5.7899 \end{pmatrix} \pm \begin{pmatrix} 0.2102 & 0.0275 & 0.0647 \\ & 0.2515 & 0.0616 \\ & & 0.2779 \end{pmatrix}$
(1000, 50)	$\begin{pmatrix} 4.0145 & 0.8003 & 0.2068 \\ & 5.0017 & -0.1926 \\ & & 6.0022 \end{pmatrix} \pm \begin{pmatrix} 0.0994 & 0.0123 & 0.0290 \\ & 0.1343 & 0.0307 \\ & & 0.1352 \end{pmatrix}$	$\begin{pmatrix} 4.0052 & 0.7996 & 0.2125 \\ & 4.9694 & -0.1884 \\ & & 5.9826 \end{pmatrix} \pm \begin{pmatrix} 0.1083 & 0.0128 & 0.0307 \\ & 0.1344 & 0.0337 \\ & & 0.1358 \end{pmatrix}$
(5000, 10)	$\begin{pmatrix} 4.0149 & 0.8004 & 0.2068 \\ & 5.0030 & -0.1925 \\ & & 6.0012 \end{pmatrix} \pm \begin{pmatrix} 0.0427 & 0.0067 & 0.0104 \\ & 0.0545 & 0.0078 \\ & & 0.0481 \end{pmatrix}$	$\begin{pmatrix} 4.0107 & 0.7984 & 0.2114 \\ & 4.9888 & -0.1903 \\ & & 6.0137 \end{pmatrix} \pm \begin{pmatrix} 0.0536 & 0.0075 & 0.0128 \\ & 0.0644 & 0.0105 \\ & & 0.0543 \end{pmatrix}$

(b)

(N, L)	$\langle \Sigma \rangle \pm \sigma_\Sigma$	$\langle \hat{\Sigma} \rangle \pm \sigma_{\hat{\Sigma}}$
(250, 200)	$10^{-4} \begin{pmatrix} 2.0158 & 10.0939 & 0.6220 \\ & 78.9217 & -3.6252 \\ & & 4.4858 \end{pmatrix} \pm 10^{-5} \begin{pmatrix} 1.7349 & 10.4692 & 1.9448 \\ & 73.8076 & 11.3638 \\ & & 3.5033 \end{pmatrix}$	$10^{-4} \begin{pmatrix} 1.7766 & 8.7370 & 0.5382 \\ & 69.3905 & -3.6824 \\ & & 4.1709 \end{pmatrix} \pm 10^{-5} \begin{pmatrix} 1.8817 & 10.7886 & 1.8881 \\ & 72.2916 & 10.9173 \\ & & 3.5979 \end{pmatrix}$
(1000, 50)	$10^{-4} \begin{pmatrix} 2.0152 & 10.0911 & 0.6218 \\ & 78.8982 & -3.6235 \\ & & 4.4849 \end{pmatrix} \pm 10^{-5} \begin{pmatrix} 0.8817 & 5.5193 & 0.8904 \\ & 40.0317 & 6.3270 \\ & & 1.7223 \end{pmatrix}$	$10^{-4} \begin{pmatrix} 1.9997 & 9.9635 & 0.6333 \\ & 77.6466 & -3.4983 \\ & & 4.4423 \end{pmatrix} \pm 10^{-5} \begin{pmatrix} 0.9640 & 5.7310 & 0.9719 \\ & 40.1822 & 6.6121 \\ & & 1.7890 \end{pmatrix}$
(5000, 10)	$10^{-4} \begin{pmatrix} 2.0156 & 10.0939 & 0.6217 \\ & 78.9069 & -3.6218 \\ & & 4.4841 \end{pmatrix} \pm 10^{-5} \begin{pmatrix} 0.4515 & 2.8824 & 0.3099 \\ & 18.8103 & 1.6360 \\ & & 0.6727 \end{pmatrix}$	$10^{-4} \begin{pmatrix} 2.0082 & 10.0135 & 0.6350 \\ & 78.3321 & -3.5694 \\ & & 4.4935 \end{pmatrix} \pm 10^{-5} \begin{pmatrix} 0.5842 & 3.5035 & 0.4149 \\ & 21.5692 & 1.9543 \\ & & 0.7955 \end{pmatrix}$

Table 2 Results using two scalar Hénon time series corrupted by high level noises.

The two scalar noisy time series are obtained from the observation, using as observables $h^1(y_1, y_2) = \frac{50y_2}{5+y_1}$ and $h^2(y_1, y_2) = 4y_1^2 + y_2$ respectively, of an orbit of the Hénon map, and corrupted by noises with noise-to-signal levels of 50% and 25%, respectively, and with a correlation of $R_{12} = 0.75$.

(a) True (on the left of the table) and estimated (on the right of the table) mean noise-to-signal levels, and mean correlations between the coordinates of the two dimensional error term obtained using L time series each of N data points. The corresponding standard deviations appear in the matrix beside the covariance matrix.

(b) True (on the left of the table) and estimated (on the right of the table) mean covariance matrices obtained using L time series each of N data points. The corresponding standard deviations appear in the matrix beside the covariance matrix.

(a)

(N, L)	$\langle \Gamma \rangle \pm \sigma_\Gamma$	$\langle \hat{\Gamma} \rangle \pm \sigma_{\hat{\Gamma}}$
(250, 200)	$\begin{pmatrix} 50.0479 & 0.7479 \\ & 24.9083 \end{pmatrix} \pm \begin{pmatrix} 2.5801 & 0.0316 \\ & 1.3589 \end{pmatrix}$	$\begin{pmatrix} 49.1758 & 0.7123 \\ & 26.7570 \end{pmatrix} \pm \begin{pmatrix} 3.2230 & 0.0430 \\ & 2.0494 \end{pmatrix}$
(1000, 50)	$\begin{pmatrix} 50.1225 & 0.7480 \\ & 24.9446 \end{pmatrix} \pm \begin{pmatrix} 1.4944 & 0.0148 \\ & 0.7209 \end{pmatrix}$	$\begin{pmatrix} 49.6568 & 0.7380 \\ & 25.440 \end{pmatrix} \pm \begin{pmatrix} 1.7531 & 0.0172 \\ & 0.9271 \end{pmatrix}$
(5000, 10)	$\begin{pmatrix} 50.1422 & 0.7480 \\ & 24.9567 \end{pmatrix} \pm \begin{pmatrix} 0.6323 & 0.0065 \\ & 0.3971 \end{pmatrix}$	$\begin{pmatrix} 49.6780 & 0.7399 \\ & 25.3792 \end{pmatrix} \pm \begin{pmatrix} 0.7206 & 0.0094 \\ & 0.3732 \end{pmatrix}$

(b)

(N, L)	$\langle \Sigma \rangle \pm \sigma_\Sigma$	$\langle \hat{\Sigma} \rangle \pm \sigma_{\hat{\Sigma}}$
(250, 200)	$\begin{pmatrix} 1.4172 & 0.4481 \\ & 0.2533 \end{pmatrix} \pm \begin{pmatrix} 0.1316 & 0.0519 \\ & 0.0238 \end{pmatrix}$	$\begin{pmatrix} 1.3463 & 0.4337 \\ & 0.2753 \end{pmatrix} \pm \begin{pmatrix} 0.1307 & 0.0554 \\ & 0.0316 \end{pmatrix}$
(1000, 50)	$\begin{pmatrix} 1.4176 & 0.4483 \\ & 0.2534 \end{pmatrix} \pm 10^{-2} \begin{pmatrix} 7.5330 & 2.8131 \\ & 1.3130 \end{pmatrix}$	$\begin{pmatrix} 1.3845 & 0.4387 \\ & 0.2552 \end{pmatrix} \pm 10^{-2} \begin{pmatrix} 7.6815 & 2.9871 \\ & 1.4727 \end{pmatrix}$
(5000, 10)	$\begin{pmatrix} 1.4174 & 0.4483 \\ & 0.2534 \end{pmatrix} \pm 10^{-2} \begin{pmatrix} 3.1851 & 1.2760 \\ & 7.1524 \end{pmatrix}$	$\begin{pmatrix} 1.3991 & 0.4429 \\ & 0.2561 \end{pmatrix} \pm 10^{-2} \begin{pmatrix} 3.4604 & 1.2740 \\ & 5.8859 \end{pmatrix}$

Table 3 Results obtained using three scalar Lorenz time series corrupted by noises of moderate amplitude and with very small correlations.

The three noisy scalar time series are obtained from the observation, using the observables $h^1(y_1, y_2, y_3) = y_1 + 5y_2^2 + y_3$, $h^2(y_1, y_2, y_3) = 2y_2 + y_3^2$, and $h^3(y_1, y_2, y_3) = 2y_1 + y_3^2$, respectively, of an orbit of the Lorenz map corrupted by noises with noise-to-signal levels of 5%, 10%, and 20%, respectively, and correlations $R_{12} = 0.1$, $R_{13} = 0$, and $R_{23} = -0.1$.

(a) True (on the left of the table) and estimated (on the right of the table) mean noise-to-signal levels, and mean correlations between the coordinates of the three-dimensional error term obtained using L time series each of N data points. The corresponding standard deviations appear in the matrix beside the covariance matrix.

(b) True (on the left of the table) and estimated (on the right of the table) mean covariance matrices of the three-dimensional error term obtained using L time series each of N data points. The corresponding standard deviations appear in the matrix beside the covariance matrix.

(a)

(N, L)	$\langle \Gamma \rangle \pm \sigma_\Gamma$	$\langle \hat{\Gamma} \rangle \pm \sigma_{\hat{\Gamma}}$
(250, 200)	$\begin{pmatrix} 4.9794 & 0.0976 & 0.0016 \\ & 10.0405 & -0.0927 \\ & & 20.0086 \end{pmatrix} \pm \begin{pmatrix} 0.4025 & 0.0657 & 0.0663 \\ & 0.7427 & 0.0641 \\ & & 1.5682 \end{pmatrix}$	$\begin{pmatrix} 4.9408 & 0.0394 & -0.0161 \\ & 9.1081 & -0.1893 \\ & & 17.4878 \end{pmatrix} \pm \begin{pmatrix} 0.5937 & 0.0913 & 0.0759 \\ & 0.8092 & 0.0832 \\ & & 1.6760 \end{pmatrix}$
(1000, 50)	$\begin{pmatrix} 4.9703 & 0.0977 & 0.0017 \\ & 10.0267 & -0.0930 \\ & & 19.9797 \end{pmatrix} \pm \begin{pmatrix} 0.2075 & 0.0395 & 0.0354 \\ & 0.4058 & 0.0333 \\ & & 0.7961 \end{pmatrix}$	$\begin{pmatrix} 5.1547 & 0.0640 & 0.0078 \\ & 9.9787 & -0.1190 \\ & & 19.2181 \end{pmatrix} \pm \begin{pmatrix} 0.3660 & 0.0551 & 0.0458 \\ & 0.4856 & 0.0495 \\ & & 0.8605 \end{pmatrix}$
(5000, 10)	$\begin{pmatrix} 4.9673 & 0.9777 & 0.0016 \\ & 10.0216 & -0.0929 \\ & & 19.9675 \end{pmatrix} \pm \begin{pmatrix} 0.1255 & 0.0125 & 0.0227 \\ & 0.1936 & 0.0124 \\ & & 0.4790 \end{pmatrix}$	$\begin{pmatrix} 4.9777 & 0.0826 & 0.0269 \\ & 10.1152 & -0.1125 \\ & & 19.4847 \end{pmatrix} \pm \begin{pmatrix} 0.1782 & 0.0172 & 0.0303 \\ & 0.1836 & 0.0137 \\ & & 0.4588 \end{pmatrix}$

(b)

(N, L)	$\langle \Sigma \rangle \pm \sigma_\Sigma$	$\langle \hat{\Sigma} \rangle \pm \sigma_{\hat{\Sigma}}$
(250, 200)	$\begin{pmatrix} 4610.3614 & 738.0409 & 23.8038 \\ & 12412.3641 & -2292.0197 \\ & & 49267.5000 \end{pmatrix} \pm \begin{pmatrix} 441.3230 & 503.2981 & 1003.5568 \\ & 1076.740 & 1591.1875 \\ & & 4472.5061 \end{pmatrix}$	$\begin{pmatrix} 4545.6943 & 268.4710 & -211.2331 \\ & 10231.6440 & -3730.2993 \\ & & 37967.8924 \end{pmatrix} \pm \begin{pmatrix} 753.5591 & 624.3761 & 1003.9602 \\ & 1117.0072 & 1601.6918 \\ & & 4310.4246 \end{pmatrix}$
(1000, 50)	$\begin{pmatrix} 4609.1656 & 738.8583 & 25.3161 \\ & 12415.1600 & -2301.3570 \\ & & 49282.1260 \end{pmatrix} \pm \begin{pmatrix} 224.1265 & 298.8531 & 535.2331 \\ & 578.4700 & 821.1563 \\ & & 2100.6895 \end{pmatrix}$	$\begin{pmatrix} 4967.3241 & 501.1333 & 117.3745 \\ & 12347.6393 & -2836.9241 \\ & & 46055.0945 \end{pmatrix} \pm \begin{pmatrix} 581.2906 & 431.3763 & 693.3772 \\ & 756.9442 & 1117.9904 \\ & & 2325.5794 \end{pmatrix}$
(5000, 10)	$\begin{pmatrix} 4608.5088 & 738.9274 & 24.4205 \\ & 12414.9072 & -2297.3404 \\ & & 49276.7200 \end{pmatrix} \pm \begin{pmatrix} 112.8640 & 100.2151 & 342.1090 \\ & 152.6486 & 293.4352 \\ & & 1036.1637 \end{pmatrix}$	$\begin{pmatrix} 4625.2875 & 632.7281 & 398.1572 \\ & 12673.2134 & -2754.8109 \\ & & 47324.8654 \end{pmatrix} \pm \begin{pmatrix} 205.0083 & 138.2098 & 446.0139 \\ & 208.6688 & 316.3352 \\ & & 974.9102 \end{pmatrix}$